Review

Affective algorithmic composition of music: A systematic review

Abigail Wiafe and Pasi Fränti*

School of Computing, University of Eastern Finland, Joensuu, Finland
* **Correspondence:** franti@cs.uef.fi

Abstract: Affective music composition systems are known to trigger emotions in humans. However, the design of such systems to stimulate users' emotions continues to be a challenge because, studies that aggregate existing literature in the domain to help advance research and knowledge is limited. This study presents a systematic literature review on affective algorithmic composition systems. Eighteen primary studies were selected from IEEE Xplore, ACM Digital Library, SpringerLink, PubMed, ScienceDirect, and Google Scholar databases following a systematic review protocol. The findings revealed that there is a lack of a unique definition that encapsulates the various types of affective algorithmic composition systems. Accordingly, a unique definition is provided. The findings also show that most affective algorithmic composition systems are designed for games to provide background music. The generative composition method was the most used compositional approach. Overall, there was rather a low amount of research in the domain. Possible reasons for these trends are the lack of a common definition for affective music composition systems and also the lack of detailed documentation of the design, implementation and evaluation of the existing systems.

Keywords: affect; affective algorithmic systems; affective music composition; musical features; music; systematic review

1. Introduction

Early investigations [1] have demonstrated that digital processing tools can generate new "unheard sounds" which can elicit emotions. These tools may use subjective music experiences of listeners and machine learning (ML) methods to create music that evokes affect [2]. This is known as *affective algorithmic composition* (AAC) of music. AAC can control listeners' moods, reduce stress, and aid meditation [3]. AAC systems have been shown to be effective in areas including games [4–6], therapeutics [7,8], and emotion recognition [9–11]. Arguably, AAC adoption is attributed to the fact that it provides an easy means for both novice and expert music creators to compose music [12].

Several studies have advanced AAC research by providing new methods [7,9] to aid its design

and implementation. However, extensive knowledge on patterns, types, definitions, algorithms, and composition methods that are pertinent for advancing the field is limited [13–16].

To address the above limitation, this paper presents a systematic review of literature on existing AAC studies. Through an extensive literature search and review, we also propose a new definition for AAC systems that encapsulates the various types of systems. This definition addresses issues regarding inconsistencies in current definitions that have been identified during the review.

Furthermore, we identify procedural composition methods as the dominant type of method used for background music in games. This method is effective for the timing of events, actions and player interactivity. Our findings also outline challenges and some future directions in the domain. Section 2 presents the review questions and details of the steps used to select and analyze the studies. This is followed by the findings and discussions in Sections 3 and 4, respectively, and the conclusions are drawn in Section 5.

2. Review approach

We adopted the approach of Kitchenham and Charters [17], which has been used in machine learning (ML) and e-learning [19–17]. It provides a rigorous and impartial assessment tool for posing review questions, selecting studies, analysis and reporting findings. After a preliminary investigation of existing literature, we identified knowledge gaps that led to the following research questions (RQs):

- RQ1: What are the characteristics of current studies?
- RQ2: How do existing studies define AAC systems?
- RQ3: What are the emotional models used for designing AAC systems?
- RQ4: What ML methods are dominant in these systems?
- RQ5: What music composition methods are used?
- RQ6: What practical applications are considered?
- RQ7: What musical features are manipulated to generate affect?
- RQ8: What methods are used for evaluating the performance of AAC systems?

RQ1: Current summaries in AAC studies provide limited information on publication patterns, dominating outlets and leading research venues. Thus, it is expected that answering this question will provide pertinent information to researchers and practitioners to facilitate new knowledge generation.
RQ2: Preliminary readings in the domain provided inconsistent and diverging definitions of AAC. Hence, responses to this question will require finding existing patterns in definitions to provide a unified definition that encapsulates the various types of AAC.
RQ3: Existing systems use emotional models for measuring affects, but they do not provide information on the most appropriate and dominant models to guide developments. Answers to this question are relevant because different systems may require different emotional models.
RQ4: ML methods provide potential opportunities for developing intelligent systems; however, there is limited evidence on patterns and ML methods that dominate the domain. Answers to this question will provide information regarding which ML methods are used for composing and detecting users' affective states.
RQ5: This question will answer which compositional methods are more dominant in AAC systems and what types of affect these systems seek to alter.
RQ6: Existing studies have outlined some practical applications; however, the explanatory and

statistical information on how they are implemented and used in various research fields is inadequate. Answers to this question will provide better understanding of how to use these systems in practical applications.

RQ7: Existing studies show that musical features are altered to provoke emotion, but information about dominant features and generated affect types are scarce. Information on the types of musical features used to target a specific affect has been discussed less often. Answering this question will provide information on how these features can be implemented and manipulated during system design.

RQ8: Existing studies have demonstrated that listeners use their experiences to assess emotions. However, literature on appropriate evaluation methods for assessing system performance is limited.

2.1. Selection criteria and the process

The ability to select relevant studies is paramount in reviews as the most relevant and timely research data are identified. Research validation is also maximized as feasibility is improved and ethical concerns are minimized. The studies also offer structured, reproducible means [17]. We used the following review protocol. The protocol consists of a set of inclusion and exclusion criteria for the study. Figure 1 is a diagrammatic representation of the processes adopted for the review.
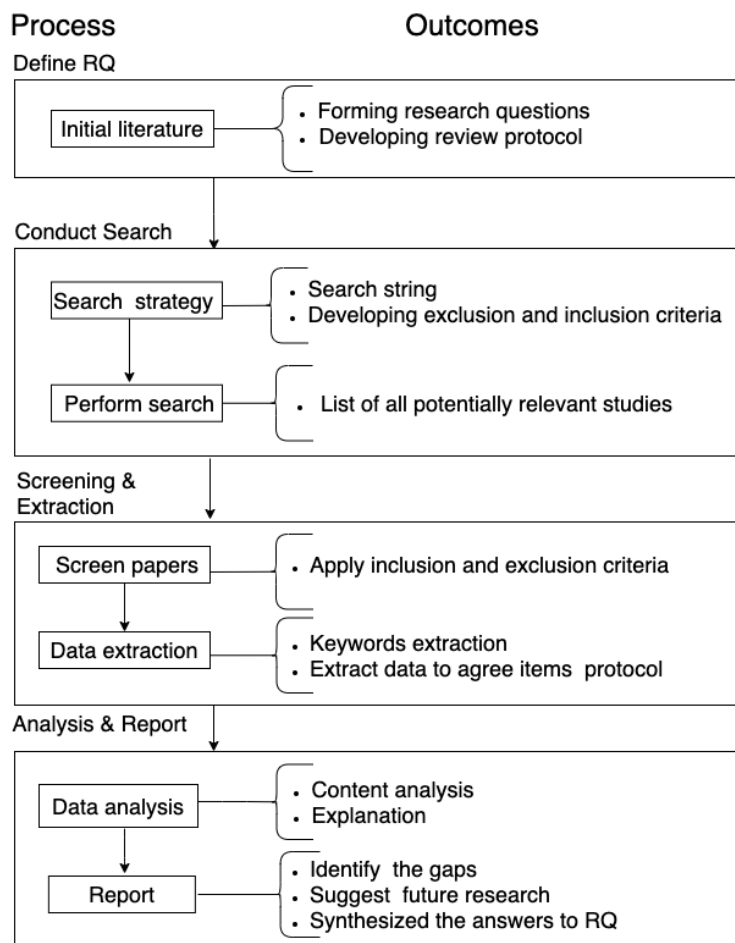

Figure 1. Main phases of the review process.

A literature search was conducted to extract peer-reviewed articles from a large body of research work. Conforming to existing systematic review processes [21,22], that have been shown to be effective [23,24], content and qualitative research techniques for analyzing text-based data were used during the preliminary stages of the study. This provided relevant information for identifying the appropriate search terms, phrases, and databases. We found the following search terms most appropriate for identifying relevant primary studies:

- *Affective music* AND *composition*
- *Affective music* AND *generation*
- *Machine learning* AND *affective music*
- *Algorithmic* AND *affective music*

We searched six databases: ACM Digital library, IEEE Xplore, PubMed, ScienceDirect, Google Scholar, and SpringerLink. In each case, the "*All text*" field of the advanced search functionality was used to filter the studies. Studies published between January 2015 and December 2020 were considered. This ensured that our findings would be recent. All publications were in English, peer-reviewed and reported in conference proceedings or journals. Book chapters, reviews, magazines, editorials and letters were excluded.
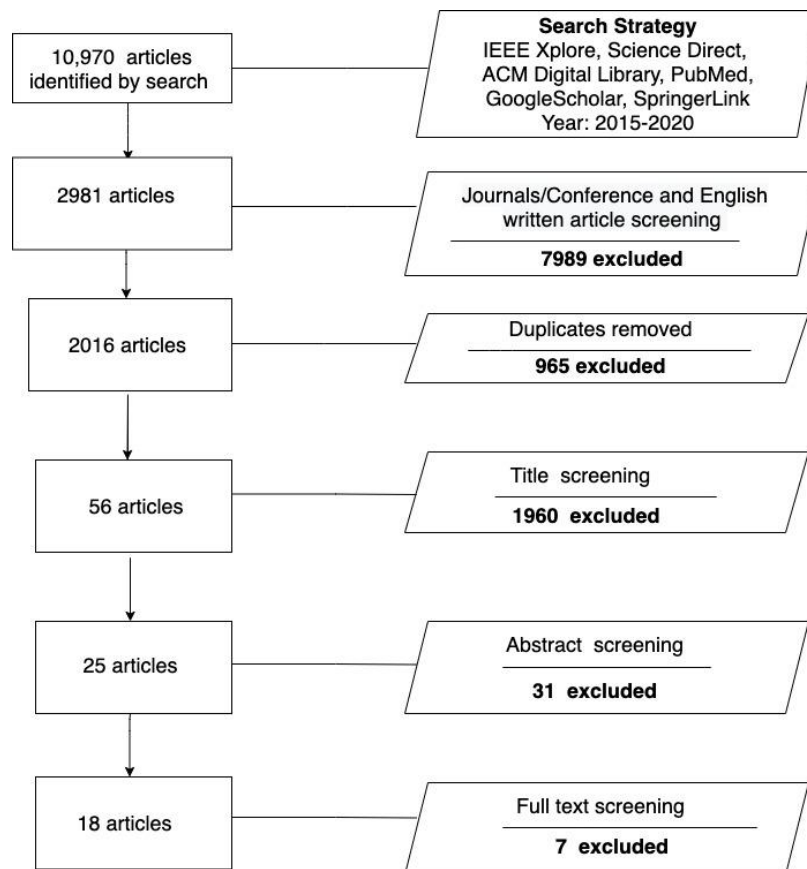

Figure 2. Overview of the screening and selection process for this study.

A review of titles and abstracts was conducted first to eliminate studies that do not suggest any form of music composition that uses an algorithm to induce a listener's mood. This reduced the number of articles from 10,970 to 2,981. The titles were then manually reviewed and 7,989 studies

were removed. Duplicates were removed. Titles were screened manually by using the inclusion criteria. Each abstract was read to select studies that suggest the development of an AAC system. Papers whose titles and abstracts matched our objectives were read in more detail. Figure 1 presents the three main activities performed and the steps used at each stage. All selected studies discuss either concepts, techniques, approaches, emotional models, musical features, or the application of affective music composition systems. After applying all the filtration rules, a total of 18 primary studies were obtained. Figure 2 is a diagrammatic representation of the stages of the study selection process.

3. Findings and results

3.1. Characteristics of the studies (RQ1)

All 18 studies discuss a form of the AAC system, although some did not explicitly mention them. The main observation is that, during the six-year period under review, research on AAC systems has not received adequate attention. The low number of studies is not encouraging. This finding is contrary to claims made by Gonzalez and McMullen [9], which suggested significant growth in AAC research.

All studies originated from eight countries (see Figure 3), and studies that were from Europe dominated (10 studies were from Europe). Five studies had their corresponding/lead authors from the United Kingdom. Studies from the United States of America, Japan, Taiwan and Australia were relatively low in number. This was not encouraging considering that these countries dominate ML research in general [25]. Seven studies were published in journals. Studies were distributed among 17 scholarly outlets, and more studies were recorded in 2015. Figure 3 provides details on the distribution of studies and their corresponding emanating countries.
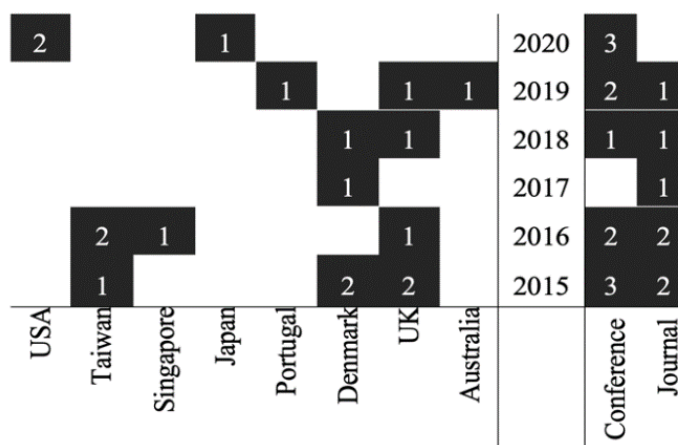

Figure 3. Study characteristics by country and publication outlets.

3.2. Defining and categorizing AAC systems (RQ2)

The description of AAC systems in the literature are ambiguous. Twelve of the primary studies did not define them, although they successfully demonstrated systems capable of composing music that alters listeners' moods. One study [S2] defined AAC systems as part of generic affective

computing technologies. This is a valid definition; however, there is still a need to distinguish AAC systems from other types of affective computing technologies. Findings from this study also suggest that there are different types of AAC systems. In some studies [S1, S3], they are defined as systems that integrate affective science and computer-aided composition for the generation of new music. This definition conforms to existing ones [13,26] since it caters to both induced and perceived emotional changes in listeners. In one study [S5], AAC is classified as a combination of *automatic accompaniment generation* and *affective music synthesis*. Thus, it is considered to generate new music and also modify existing melodies to target affect. Another study [S7] explained how it inherits perceptual models in generic computer science without emphasizing the need for it to target affect. In summary, current studies define AAC as *affective music composition* (i.e., music that induces affect but does not always use algorithms) or *adaptive music composition* (i.e., music that changes according to the context of the listener).

The diversity in the definitions demonstrates different views and concepts, yet there is a need for coherence. A careful observation of the various definitions shows that researchers define AAC according to the type of investigations they are performing. Specifically, the lack of a unique definition that encapsulates all types of AAC systems has encouraged researchers to consider the types of systems as its definition.

A key requirement in all AAC systems is to induce affect by composing music using algorithms. Arguably, no music is neutral. A listener's mood is almost always altered. Hence, definitions that suggest the use of music to induce affect are insufficient, as are those that focus only on the use of algorithms for music composition. The distinction between AAC systems and other algorithmic music composition systems is that AAC systems seek to deliberately *alter* or *emphasize* listeners' moods. This is not the case for algorithmic or generic music compositions. In this case, the level of alteration varies according to the selection of musical features used for composition when compared to generic music composition. For instance, one may target a single affective descriptor, a combination of descriptors or a position on a dimensional mood. One can compose unique and customized music for each listener, which is a situation that is not common in generic music composition, as it would be extremely expensive. Based on the above discussion, we provide the following definition:

> **Definition**: *Affective algorithmic composition (AAC) systems are a technology that use algorithms to compose music with the intent to modulate or induce a listener's affective state.*

According to this definition, such systems seek to provide two possible affective outcomes: *modulation* or *inducement*. First, they modulate mood changes and maintain (arousal), or vary, the magnitude of the current mood of a listener (valence). For instance, a system may reduce the intensity of sadness or joy to a particular magnitude. It may also reinforce the current mood of a listener. Second, they seek to completely change the current mood, with or without prior knowledge of the listener's current mood. For example, the system may compose music that makes listeners happy. AAC systems that seek to modulate mood must first identify the current mood of the listener. This is different for systems that seek to induce mood change. See Table 1 for examples of types of systems identified in the study.

Thirteen studies were designed to modulate the affective state of listeners. However, there was no common pattern in the targeted affective outcome, the emotional model used for composition, the design method or the area of application. It is emphasized that, although these systems produce affective music, and thus adopt concepts from affective music composition, they are not the same. AAC is rather an intersection of affective music composition and algorithmic music composition. Also, in all circumstances, the resultant affective impact on a listener (induced or modulated) must not be a side effect of using the system or listening to the music produced by the system. As indicated in our definition, these systems have an intent in all cases.

3.3. Emotional models in AAC systems (RQ3)

We identified two main categories of emotional models: categorical and dimensional. Categorical models are also known as discrete models, and they are characterized by words or adjectival phrases such as happy, sad, fear and anger [27] to classify emotions. Dimensional models represent emotions by using a two-dimensional valence-arousal space [28]. They recognize the ambiguity of adjectives and define emotions in terms of arousal (how exciting/calming) and valence (positive/negative); thus, emotions are mapped on a two-dimensional plane.

Dimensional and categorical emotional models aid the expression of specific emotional states. They communicate by using adjectival phrases or numerical values to indicate how emotions are represented and interpreted by humans [29]. From the selected studies, 14 used dimensional models and only one used a categorical emotional model. The others did not report the model type they used.

Table 1. Emotional models, affective outcomes, ML techniques, composition approaches and uses of AAC systems.

Affective Outcome	Emotional Model	ML Technique	Composition Approach	Area of Application
Modulate (S1, S4, S5, S6, S8, S9, S10, S11, S12, S13, S14, S15, S16)	Dimensional (S1, S2, S3, S5, S6, S7, S9, S10, S11, S12, S15, S16, S17, S18)	Supervised learning (S1, S3, S9, S17)	Transformative (S3, S5, S7, S16)	Therapeutic (S1, S9)
Induce (S2, S3, S7, S17, S18)	Categorial (S13)	Unsupervised learning (S15, S18)	Generative (S4, S6, S8, S9, S10, S11, S12, S13, S14, S15)	Games/ Movies (S4, S5, S6, S8, S10, S11, S13, S14)
	Not specified (S4, S8, S14)	Reinforcement (S2, S4, S5, S6, S7, S8, S10, S11, S12, S13, S14, S16)	Not specified (S1, S2, S17, S18)	Emotion recognition (S2, S7, S12, S15)
				Not specified (S3, S16, S17, S18)

As a dominant model used in AAC systems, dimensional models explain the valence and arousal dimensions associated with a particular emotional feeling. The valence dimension describes the pleasantness or unpleasantness of a feeling whereas the arousal dimension explains the intensity by describing the affect. Categorical emotional models assess emotions using Likert scales to capture a person's current emotional state. Some studies [29,30] have argued that categorical models are preferred in assessing emotions in general; however, our findings contradict this. Vice versa, it is the categorical models that we observed to be less preferred in AAC system design. Rather, our findings support the arguments of Brattico and Pearce [31] and Russell [28], who suggested that dimensional models are more appropriate for emotion-related studies in music.

3.4. Machine learning methods in AAC systems (RQ4)

ML methods and techniques play a key role in AAC system design. Thus, research that applies ML in music modeling and creation usually suggests model architecture, training methods and datasets. They also provide support for estimating system performance by using measures such as the sequence of likelihood.

From the primary studies reviewed, seven unique ML techniques were identified: support vector machines (SVMs), Monte Carlo, artificial neural network (ANNs), heuristics, evolutionary algorithms, Gaussian mixture modeling and dynamic programming. These methods were used for various purposes. See Table 1 and Figure 4 for a list of ML methods identified. The methods can be broadly categorized into supervised, unsupervised and reinforcement learning.
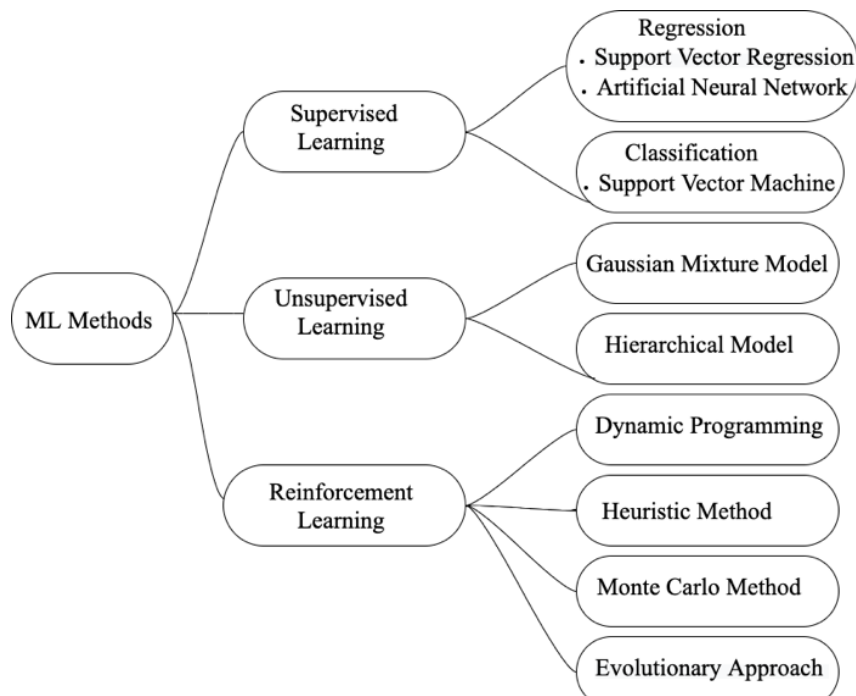

Figure 4. Types of ML methods identified in the study.

3.4.1. Supervised learning

Supervised learning techniques are used for the classification, pattern recognition and prediction of emotions [32]. Two studies used support vector machines [S1 and S9]. ANNs and support vector regression were also used [S3 and S17], respectively. ANNs are self-learning, adaptive and have no restrictions on the number of input variables; thus, they are appropriate for constructing AAC systems. Sound is mapped to parameters with labels representing qualities of musical features.

SVMs are mainly used for classifications and determinations of different emotional states. Here, the classification problem is considered a binary problem (high vs low arousal, high vs low valence), and this eliminates the possibility of measuring affect as a continuous function. This is a limitation since the two states (valence and arousal) are numerical and have more than two possible values.

3.4.2. Unsupervised learning

Two studies [S15, S18] used unsupervised learning approaches. One study [S15] used the Gaussian mixture model to compose a piece of generative music (see Table 1), whereas the other used hierarchical linear modeling (HLM) [S18]. HLM is more appropriate when data are hierarchical or clustered. It groups data into clusters and analyzes them by using statistical models to identify their effect on data. Based on similarities in characteristics (subjective and physiological), it establishes a relationship to identify how music generated correlates with specific musical features.

3.4.3. Reinforcement learning

Twelve studies used reinforcement learning methods (see Table 1). This method considers decision processes, recommendations and reward systems by learning to react to its environment. Specifically, a multi-modal and multi-agent composition method, such as the heuristic method, was used in these studies [S8, S13]. Markov chains and Monte Carlo methods were used in five studies. Three of them [S2, S4, S7] used Markov Chains, and two [S11 and S12] used stochastic processes.

The reason for Monte Carlo being more popular may be attributed to its ease of use since it makes use of probability distributions for compositions that exhibit inherent uncertainty. Different sets of random note values are calculated by using a probability function to produce music that is flexible and scalable. Features such as chords are selected and combined with rhythmic patterns (note value and time signature) and pitches to generate accompaniment. They can determine harmonic progressions as a continuous stream of chords by using random selection and iterative processes to compute successive chords. This makes it easier for the composition. Although they are preferred, some researchers [33] have argued that these methods produce low-quality and artificial music.

Three studies [S6, S10 and S14] used evolutionary methods. These methods apply optimization techniques that aim to identify the optimal music based on a given criterion. They facilitate agility in music composition because they assess the fitness of musical features (individuals) by utilizing a specified composition rule to improve music quality. All three evolutionary approaches used the feasible/infeasible two-population approach (FI-2POP) with multiple objective optimizations.

Dynamic programming (DP) was used in two studies [S5, S16], and it is among the least popular ML methods in AAC systems. In DP, music is generated by categorizing similar notes into subgroups and applying an incremental method in the composition process. DP applies sequential matching of musical notes and affects. However, this process may lead to note repetitions that result

in sound distortion. Heuristic methods were used in two studies [S8, S13]; mostly, they use multi-agent and multi-modal methods for inducing affect.

3.5. Music composition methods (RQ5)

Two main composition methods are predominant in AAC systems: generative and transformative compositions. This finding corroborates the suggestions by Wooller et al. [34] that generative and transformative approaches are the main methods used in algorithmic composition, contrary to claims in [35] which suggest the "sequence" approach. Considering that the review focused on AAC systems, there is a possibility that sequencing is used in other algorithmic music compositions but not in AACs.

The findings in Table 1 indicate that the generative approach is dominant. Generative composition produces different and unique music from scratch, whereas transformative compositions compose unique music by using existing inputs to generate music. Thus, in transformative approaches, the musical information is altered. The generative approach is used extensively for composing themes or background music in games, and also for therapeutic purposes. Nine of the generative approaches used procedural methods, and one used the structural method. Four studies used the transformative approach, while four did not specify the approach used.

Regarding studies that adopted the transformative approach (composing music using existing inputs to form a variant of the original music), it was not possible to identify a specific dominant transformation method because there were only three methods, of which two used the transposition method, and the third used the retrograde and inversion methods. The retrograde transformation method reverses notes in a musical sequence, while the transposition method moves a group of notes (pitches) up and down at a constant pitch interval. The inversion method applies changes to intervals, melodies, chords and tones of existing music to form new music. Considering that transformative approaches were less explored, there is a need for further studies to investigate their minimal usage.

3.6. Practical applications (RQ6)

We observed that AAC systems are predominantly used for composing background music (eight), and used less frequently for therapeutic purposes (two studies). Out of the eight that applied AAC in games [S4, S5, S6, S8, S10, S11, S13 and S14], seven used generative approaches, and one used the transformative approach. All eight studies sought to modulate an affective state, i.e., emphasized, increased or decreased a behavior. In games, background music is used to induce immersion and stimulate users. For instance, two studies [S10, S11] used AAC to generate real-time background music that could express specific moods.

Four of the studies focused on generic emotional recognition research, and four did not specify the domain of the application. Two studies used it for therapeutic purposes. Again, as discussed above, generative composition promotes conducive music in such instances. See Figure 5 for details.
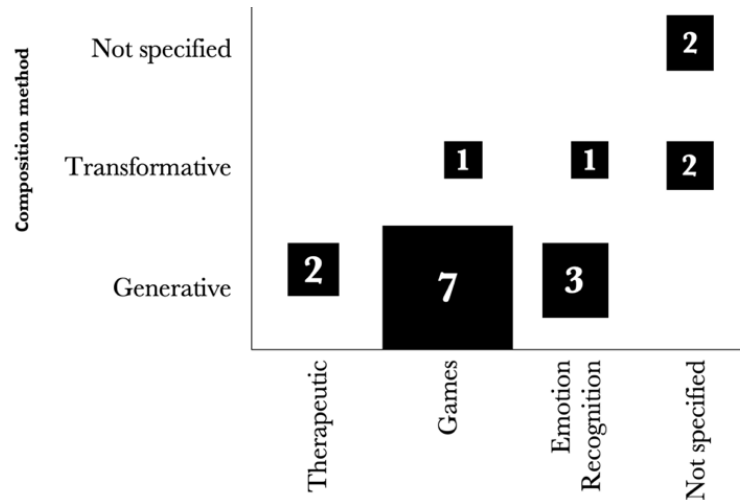
Figure 5. Composition methods and areas of application.

3.7. Musical feature manipulation (RQ7)

Cogitating that acoustic signals encrypted into auditory signals are mapped to relevant features and analyzed to convey meanings in music, there is a need to understand the types of features predominantly manipulated. Musical features hold different qualities which describe a piece of music. This quality is responsible for generating affect in listeners. Our findings suggest that the main musical features used in AAC systems are as follows:

- Harmony (11): S4, S5, S8, S10, S11, S12, S13, S14, S15, S16, S18
- Rhythm (10): S4, S5, S7, S10, S11, S12, S14, S16, S17, S18
- Dynamics (9): S3, S5, S10, S11, S13, S14, S16, S17, S18
- Timbre (7): S3, S8, S10, S11, S14, S15, S18
- Melody (7): S3, S4, S5, S8, S12, S16, S17
- Mode (6): S2, S3, S8, S16, S17, S18
- Tempo (5): S3, S5, S8, S13, S17
- Articulation (1): S13

The results indicate that harmony (11) and rhythm (10) are the two most common musical features manipulated. Articulation (one) is the least used feature considering that only one study used it. The majority of the studies that reported the type of musical feature manipulated used more than one feature. Only one study [S2] manipulated a single feature (modality). The other six features were used in at least five studies. Three studies did not report the type of features that were manipulated.

The dominance of harmony may be attributed to its ability to allow listeners to relate to music. Harmony contains pleasing sounds that are visceral or intuitive to emotional response, and it functions as the building block of chords and song structure. Likewise, rhythm provides the compositional structure, measures movement and provides motion to melody and harmony, and it can generate emotions. As the second most used feature, our finding supports Williams' [36] claims that rhythm is one of the most universally accepted features used in AAC systems.

It was also observed that there is a lack of explicit discussions on the relationship between music features, affective states and their corresponding physiological response. Particularly, patterns and

trends in the primary studies did not identify specific musical features that promote specific affects or emotions. This lack of reliable musical/audio features is concerning. See Table 2 for the distribution of musical features manipulated to generate affect in AAC systems.

Table 2. Emotional models, affective outcomes, ML techniques, composition approaches and uses of AAC systems.

Study	Music Feature	Evaluation Method	Success
S1	---	Physiological/ Self-Reporting	Yes
S2	Mode	Physiological/ Self-Reporting	Yes
S3	Tempo, mode, melody, timbre, and dynamics	Physiological/ Self-Reporting	Yes
S4	Harmony, rhythm, and melody	Self-Reporting	Yes
S5	Tempo, rhythm, dynamics, melody, and harmony	Self-Reporting	Yes
S6	---	Self-Reporting	Yes
S7	Rhythm	Self-Reporting	Yes
S8	Tempo, melody, harmony, mode, and timbre	Self-Reporting	Yes
S9	---	Physiological/ Self-Reporting	Yes
S10	Dynamics, timbre, rhythm, and harmony	Self-Reporting	Yes
S11	Dynamics, timbre, rhythm, and harmony	Self-Reporting	Yes
S12	Harmony, tempo, melody, articulation, and dynamics	Self-Reporting	Yes
S13	Harmony, melody, and rhythm	Self-Reporting	Yes
S14	Dynamics, timbre, rhythm, and harmony	---	---
S15	Timbre and harmony	Self-Reporting	Yes
S16	Melody, mode, harmony, rhythm, and dynamics	Self-Reporting	Yes
S17	Tempo, rhythm, dynamics, melody, and mode	Physiological/ Self-Reporting	Yes
S18	Dynamics, timbre, rhythm, mode, and harmony	Physiological/ Self-Reporting	Yes

3.8. Performance evaluation (RQ8)

Affective experience is an expressive charm that people encounter when listening to music. People's emotional preferences for music are notable and different due to physiological, physical and environmental factors. To assess the presence of emotions, self-reporting, observation, and physiological assessment methods are used. According to the study, self-reporting methods, including self-assessment manikin [37], FEELTRACE [38] and questionnaires, are the most commonly used type of method for validating the presence of emotion. These findings align with the results from Eerola and Vuoskoski [14], which suggested that self-reporting methods are dominant in recognizing musical emotions. Also, it was observed that all other methods use self-reporting as a complementary method in verifying emotional experience.

None of the primary studies stated how subjective questionnaires were applied to measure participants' feelings, attitudes and sensitivity. Although self-reporting is most popular due to its ease of use, the use of questionnaires may lead to inaccurate responses and dishonesty. Self-reported answers may be exaggerated and biased, which may prevent self-reporting from being the most effective evaluation method. To compensate for this shortcoming, some studies [S1, S2, S3, S9, S17 and S18] used both physiological assessment and self-reporting. None of the studies relied solely on physiological assessment. Although physiological methods are appropriate, they are not the most reliable and accurate according to Trochidis and Lui [8]. Four of the seven studies that conducted physiological assessments used brain-computer interfaces [S1, S2, S3 and S9].

4. Discussion

This study constituted a systematic review of AAC systems. All previous reviews [10,11,13,36] have focused on the use of algorithms for composing music, whereas our study provides a broader and multi-perspective overview of the current state of AAC system design. Our findings confirm existing claims [13] that research in AAC systems is dormant. Thus, it is important to emphasize the need for studies to focus on the development of theoretically sound explanations of AAC system design practices.

From our investigations, the most predominant and resounding issue is that existing literature does not provide adequate information on how AAC systems are designed. Only one study [S10] provided a detailed discussion of its system's architecture and implementation. It is essential to note that these systems are complex and multi-disciplinary. Hence, designers and researchers need to be familiar with ML techniques, emotional psychology and music composition. Accordingly, the lack of literature that provides amalgamated information from these diverse disciplines, in a simple but elaborate manner, serves as a disincentive for novice researchers. There is a need for researchers to provide relevant details on how AAC systems are designed, constructed and evaluated.

It is also intriguing to note that none of the studies considered the design context. The affective response of music is context-dependent and the user's specific response to particular music may depend on his or her current emotional state, environment and ambiance, culture, personality, health and other physiological needs. Some studies [39] have provided evidence to support arguments to favor a relationship between music and physiological needs. Yet, these issues have been largely ignored in existing research. The inclusion of the context (user and use context) is expected to enhance the user's affective experience.

The low research output in the field may be attributed to the lack of studies that provide pertinent information on composition methods, techniques and musical features. This is because AAC is relatively new and interdisciplinary, and thus may not be appealing to novice researchers. It is therefore recommended that studies that seek to design AAC should provide a detailed description of how the system was designed. This will provide clarity and encourage research in the domain.

5. Conclusions

We have presented a comprehensive review of the current state of research on AAC system designs. To the best of our knowledge, it is the first systematic review performed in this area. We followed a review protocol and provided a set of inclusion and exclusion criteria. The findings indicate diverging definitions for AAC systems; hence, a universal definition was provided in this paper. The findings suggest that harmony and rhythm are the two most common musical features manipulated, while musical features such as structure have not been used in any study.

Dimensional models were the dominant emotional models observed. Also, our findings suggest that the application of AAC in games has been promising, as a significant amount of research is conducted in gaming. The scarcity of literature in the domain serves as a challenge for novice researchers. Based on these findings, future research in the domain must endeavor to present AAC systems clearly and concisely to enable novice researchers to replicate their studies. This will encourage and promote studies on AAC.

Conflict of interest

All authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. D. M. Butler, *An Historical Investigation and Bibligraphy of Ninetheeth Century Music Psychology Literature*, Thesis, 1973.
2. C. Roads, *The Computer Music Tutorial*, MIT Press, 1996.
3. M. Scirea, P. Eklund, J. Togelius, S. Risi, Can you feel it? Evaluation of affective expression in music generated by MetaCompose, *GECCO 2017 - Proc. 2017 Genet. Evol. Comput. Conf.*, (2017), 211–218. https://doi.org/10.1145/3071178.3071314
4. M. Scirea, J. Togelius, P. Eklund, S. Risi, Towards an experiment on perception of affective music generation using MetaCompose, *GECCO 2018 Companion - Proc. 2018 Genet. Evol. Comput. Conf. Companion*, (2018), 131–132. https://doi.org/10.1145/3205651.3205745
5. G. R. Marcos, An investigation on the automatic generation of music and its application into video games, *2019 8th Int. Conf. Affect. Comput. Intell. Interact. Work. Demos, ACIIW 2019*, (2019), 21–25. https://doi.org/10.1109/ACIIW.2019.8925275
6. M. Scirea, M. J. Nelson, J. Togelius, Moody Music Generator: Characterising Control Parameters Using Crowdsourcing, *International Conference on Evolutionary and Biologically Inspired Music and Art*, (2015), 200–211. https://doi.org/10.1007/978-3-319-16498-4
7. I. Daly, D. Williams, A. Malik, J. Weaver, A. Kirke, F. Hwang, et al., Personalised, Multi-Modal, Affective State Detection for Hybrid Brain-Computer Music Interfacing, *IEEE T. Affect. Comput.*, **11** (2018), 111–124. https://doi.org/10.1109/TAFFC.2018.2801811
8. K. Trochidis, S. Lui, Modeling affective responses to music using audio signal analysis and physiology, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, **9617** (2016), 346–357. https://doi.org/10.1007/978-3-319-46282-0_22
9. E. J. S. Gonzalez, K. McMullen, The Design of an Algorithmic Modal Music Platform for Eliciting and Detecting Emotion, *8th Int. Winter Conf. Brain-Computer Interface, BCI 2020*, (2020), 31–33. https://doi.org/10.1109/BCI48061.2020.9061664
10. D. Williams, A. Kirke, E. Miranda, I. Daly, J. Hallowell, J. Weaver, J. et al., Investigating perceived emotional correlates of rhythmic density in algorithmic music composition, *ACM T. Appl. Percept.*, **12** (2015), 1–21. https://doi.org/10.1145/2749466
11. J. C. Wang, Y. H. Yang, H. M. Wang, S. K. Jeng, Modeling the affective content of music with a Gaussian mixture model, *IEEE T. Affect. Comput.*, **6** (2015), 56–68. https://doi.org/10.1109/TAFFC.2015.2397457
12. A. Chamberlain, M. Bødker, M. Kallionpää, R. Ramchurn, H. P. Gasselseder, The Design of Future Music Technologies, *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion - AM'18*, (2018), 1–2. https://doi.org/10.1145/3243274.3243314
13. D. Williams, A. Kirke, E. R. Miranda, E. Roesch, I. Daly, S. Nasuto, Investigating affect in algorithmic composition systems, *Psychol. Music*, **43** (2015) 831–854. https://doi.org/10.1177/0305735614543282
14. T. Eerola, J. K. Vuoskoski, A review of music and emotion studies: Approaches, emotion models, and stimuli, *Music Percept.*, **30** (2013), 307–340. https://doi.org/10.1525/mp.2012.30.3.307

15. D. J. Fern ández, F. Vico, AI Methods in Algorithmic Composition: A Comprehensive Survey, 2013. Accessed: Feb. 14, 2020. Available from: http://www.flexatone.net/algoNet/.

16. O. Lopez-Rincon, O. Starostenko, G. A.-S. Mart ín, Algoritmic music composition based on artificial intelligence: A survey, *2018 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, (2018), 187–193. https://doi.org/10.1109/CONIELECOMP.2018.8327197

17. B. Kitchenham, S. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, 2007, Accessed: Feb. 15, 2020. Available from: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471.

18. M. Cerqueira, P. Silva, S. Fernandes, Systematic Literature Review on Machine Learning in Software Engineering, *Am. Acad. Sci. Res. J. Eng. Technol. Sci.*, **85** (2022), 370–396.

19. M. N. Giannakos, P. Mikalef, I. O. Pappas, Systematic Literature Review of E-Learning Capabilities to Enhance Organizational Learning, *Inf. Syst. Front.*, **24** (2021), 619–635. https://doi.org/10.1007/s10796-020-10097-2

20. R. van Dinter, B. Tekinerdogan, C. Catal, Automation of systematic literature reviews: A systematic literature review, *Inf. Softw. Technol.*, **136** (2021), 106589. https://doi.org/10.1016/j.infsof.2021.106589

21. L. M. Kmet, R. C. Lee, L. S. Cook, *Standard quality assessment criteria for evaluating primary research papers from a variety of fields,* 2004.

22. M. Schreier, *Qualitative content analysis in practice*, Sage Publications, 2012.

23. H. Koechlin, R. Coakley, N. Schechter, C. Werner, J. Kossowsky, The role of emotion regulation in chronic pain: A systematic literature review, *J. Psychosom. Res.*, **107** (2018), 38–45. https://doi.org/10.1016/j.jpsychores.2018.02.002

24. T. Materla, E. A. Cudney, J. Antony, The application of Kano model in the healthcare industry: a systematic literature review, *Total Qual. Manag. Bus. Excell.*, **30** (2019), 660–681. https://doi.org/10.1080/14783363.2017.1328980

25. D. Ni, Z. Xiao, M. K. Lim, A systematic review of the research trends of machine learning in supply chain management, *Int. J. Mach. Learn. Cybern.*, **11** (2020), 1463–1482. https://doi.org/10.1007/s13042-019-01050-0

26. A. Mattek, *Emotional Communication in Computer Generated Music: Experimenting with Affective Algorithms*, 2011.

27. Y. Feng, Y. Zhuang, Y. Pan, Music information retrieval by detecting mood via computational media aesthetics, *Proceedings - IEEE/WIC International Conference on Web Intelligence, WI 2003*, (2003), 235–241. https://doi.org/10.1109/WI.2003.1241199

28. J. A. Russell, A circumplex model of affect, *J. Pers. Soc. Psychol.*, **39** (1980), 1161–1178. https://doi.org/10.1037/h0077714

29. T. Eerola, J. K. Vuoskoski, A comparison of the discrete and dimensional models of emotion in music, *Psychol. Music*, **39** (2011), 18–49. https://doi.org/10.1177/0305735610362821

30. R. A. Calvo, S. Mac Kim, Emotions in text: Dimensional and categorical models, *Comput. Intell.*, **29** (2013), 527–543. https://doi.org/10.1111/j.1467-8640.2012.00456.x

31. E. Brattico, M. Pearce, The neuroaesthetics of music, *Psychol. Aesthetics, Creat. Arts*, **7** (2013), 48–61. https://doi.org/10.1037/a0031624

32. S. Cunningham, H. Ridley, J. Weinel, R. Picking, Supervised machine learning for audio

emotion recognition: Enhancing film sound design using audio features, regression models and artificial neural networks, *Pers. Ubiquitous Comput.*, **25** (2021), 637–650. https://doi.org/10.1007/s00779-020-01389-0

33. R. L. De Mantaras, Making Music with AI: Some examples, *Proceeding of the 2006 conference on Rob Milne: A Tribute to a Pioneering AI Scientist, Entrepreneur and Mountaineer*, (2006), 90–100. Available from: http://portal.acm.org/citation.cfm?id=1565089

34. R. Wooller, A. Brown, E. Miranda, J. Diederich, A framework for comparison of process in algorithmic music systems, *Generative Arts Practice*, (2005), 109–124.

35. R. Rowe, *Interactive music systems: machine listening and composing*. Cambridge, Mass.: MIT Press, 1992. https://doi.org/10.2307/3680494

36. D. Williams, A. Kirke, E. R. Miranda, E. B. Roesch, S. J. Nasuto, Towards Affective Algorithmic Composition, *The 3rd International Conference on Music & Emotion, Jyväskylä, Finland, June 11-15, 2013,* 2013.

37. M. M. Bradley, P. J. Lang, Measuring emotion: The self-assessment manikin and the semantic differential, *J. Behav. Ther. Exp. Psychiatry*, **25** (1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

38. R. Cowie, E. Douglas-Cowie, S. Savvidou, E. Mcmahon, M. Sawey, M. Schröder, 'Feeltrace': An instrument for recording perceived emotion in real time, *ISCA Work. Speech & Emot.*, (2000), 19–24.

39. P. Evans, G. E. McPherson, J. W. Davidson, The role of psychological needs in ceasing music and music learning activities, *Psychol. Music*, **41** (2013), 600–619. https://doi.org/10.1177/0305735612441736

Appendix: Literature selected for the review

ID	Details
S1	I. Daly et al., "Personalised, Multi-Modal, Affective State Detection for Hybrid Brain-Computer Music Interfacing," IEEE Trans. Affect. Comput., vol. 11, no. 1, pp. 111–124, 2020, DOI: 10.1109/TAFFC.2018.2801811.
S2	E. J. S. Gonzalez and K. McMullen, "The Design of an Algorithmic Modal Music Platform for Eliciting and Detecting Emotion," 8th Int. Winter Conf. Brain-Computer Interface, BCI 2020, pp. 31–33, 2020, DOI: 10.1109/BCI48061.2020.9061664.
S3	I. Daly et al., "Towards human-computer music interaction: Evaluation of an affectively-driven music generator via galvanic skin response measures," 2015 7th Comput. Sci. Electron. Eng. Conf. CEEC 2015 - Conf. Proc., pp. 87–92, 2015, DOI: 10.1109/CEEC.2015.7332705.
S4	G. R. Marcos, "An investigation on the automatic generation of music and its application into video games," 2019 8th Int. Conf. Affect. Comput. Intell. Interact. Work. Demos, ACIIW 2019, pp. 21–25, 2019, doi: 10.1109/ACIIW.2019.8925275.
S5	Y. C. Wu and H. H. Chen, "Generation of Affective Accompaniment in Accordance with Emotion Flow," IEEE/ACM Trans. Audio Speech Lang. Process., vol. 24, no. 12, pp. 2277–2287, 2016, doi: 10.1109/TASLP.2016.2603006.
S6	M. Scirea, J. Togelius, P. Eklund, and S. Risi, "Towards an experiment on perception of affective music generation using MetaCompose," GECCO 2018 Companion - Proc. 2018 Genet. Evol. Comput. Conf. Companion, pp. 131–132, 2018, doi: 10.1145/3205651.3205745.

ID	Details
S7	D. Williams et al., "Investigating perceived emotional correlates of rhythmic density in algorithmic music composition," ACM Trans. Appl. Percept., vol. 12, no. 3, 2015, doi: 10.1145/2749466.
S8	M. Washburn and F. Khosmood, "Dynamic Procedural Music Generation from NPC Attributes," ACM Int. Conf. Proceeding Ser., 2020, doi: 10.1145/3402942.3409785.
S9	I. Daly et al., "Affective brain-computer music interfacing," J. Neural Eng., vol. 13, no. 4, pp. 1–14, 2016, doi: 10.1088/1741-2560/13/4/046022.
S10	M. Scirea, J. Togelius, P. Eklund, and S. Risi, "Affective evolutionary music composition with MetaCompose," Genet. Program. Evolvable Mach., vol. 18, no. 4, pp. 433–465, 2017, doi: 10.1007/s10710-017-9307-y.
S11	M. Scirea, M. J. Nelson, and J. Togelius, "Moody Music Generator: Characterising Control Parameters Using Crowdsourcing," in 4th International Conference, EvoMUSART 2015, Copenhagen, Denmark, April 8-10, 2015, Proceedings, 2015, vol. 9027, pp. 200–211, doi: 10.1007/978-3-319-16498-4.
S12	M. Seiça, A. Rodrigues, F. Am, and P. Machado, "Computer Generation and Perception Evaluation of Music-Emotion Associations," 14th Int. Symp. Comput. Music Multidiscip. Res., pp. 789–800, 2019.
S13	P. E. Hutchings and J. McCormack, "Adaptive Music Composition for Games," IEEE Trans. Games, vol. 12, no. 3, pp. 270–280, 2020, doi: 10.1109/TG.2019.2921979.
S14	M. Scirea, P. Eklund, and J. Togelius, "Toward a Context Sensitive Music Generator for Affective State Expression," 2015, no. June, pp. 2–3.
S15	J. C. Wang, Y. H. Yang, H. M. Wang, and S. K. Jeng, "Modeling the affective content of music with a Gaussian mixture model," IEEE Trans. Affect. Comput., vol. 6, no. 1, pp. 56–68, 2015, doi: 10.1109/TAFFC.2015.2397457.
S16	Y.-C. Wu and H. Chen, "Emotion - flow guided music accompaniment generation," in EEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2016.
S17	K. Miyamoto, H. Tanaka, and S. Nakamura, "Music generation and emotion estimation from EEG signals for inducing affective states," ICMI 2020 Companion - Companion Publ. 2020 Int. Conf. Multimodal Interact., pp. 487–491, 2020, doi: 10.1145/3395035.3425225.
S18	K. Trochidis and S. Lui, "Modeling affective responses to music using audio signal analysis and physiology," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9617 LNCS, pp. 346–357, 2016, doi: 10.1007/978-3-319-46282-0_22.